09-08-00

A

UTILITY PATENT APPLICATION TRANSMITTAL *(Only for new nonprovisional application under 37 C.F.R § 1.53(b))*	Docket No. **UDN0003** First Inventor or Application Identifier: **Shankar Iyer, et al.** Title: **METHOD AND APPARATUS FOR DETERMINING LATENCY BETWEEN MULTIPLE SERVERS AND A CLIENT** Express Mail Label No. **EL696097234US**

CERTIFICATE OF EXPRESS MAILING
"EXPRESS MAIL" Mailing Label No. **EL696097234US** Date of Deposit: **September 7, 2000**
I hereby certify that this trademark application and the papers listed thereon are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR. 1.10 on the date indicated above, and is addressed to: Box Patent Application, Assistant Commissioner for Patents, Washington, D.C. 20231

jc930 U.S. PTO
09/07/00

Name of person mailing: Amy Alba
Signature of person mailing: ____________

TO THE ASSISTANT COMMISSIONER FOR PATENTS
Box Patent Application
Washington, D.C. 20231

jc564 U.S. PTO
09/657016
09/07/00

Invention Title: **METHOD AND APPARATUS FOR DETERMINING LATENCY BETWEEN MULTIPLE SERVERS AND A CLIENT**

Inventors: **SHANKAR IYER and SRIDHARA LANKA**

If a CONTINUATION APPLICATION, *check appropriate box and supply the requisite information*

[ ] Continuation [ ] Divisional [ ] Continuation-in-part (CIP) of prior application no.: ____________

Enclosed are:

APPLICATION ELEMENTS

1. **[X]** Filing Fee as calculated and transmitted as described below
2. **[X]** Specification having **18** pages and including the following:
   - Descriptive title of the Invention
   - Cross References to Related Applications *(if applicable)*
   - Statement Regarding Federally-sponsored Research/Development *(if applicable)*
   - Reference to Microfiche Appendix *(if applicable)*
   - Background of the Invention
   - Brief Summary of the Invention
   - Brief Description of Drawings (*if drawings filed)*
   - Detailed Description
   - Claims as Classified Below
   - Abstract of the Disclosure
3. **[X]** Drawings (*when necessary as prescribed by 35 USC 113)* **Number or Sheets: 6 (Figs 1-6)**
4. Oath or Declaration
   [X] Newly executed (original)
5. [ ] Microfiche Computer Program
6. [ ] Nucleotide and/or Amino Acid Sequence Submission

ACCOMPANYING PARTS

7. **[X] Assignment Papers** (coversheet and document(s))
8. **[X] 37 C.F.R. §(b) Statement** (when there is an assignee) [X] Power of Attorney
9. [ ] English Translation Document (if applicable)
10. [ ] Information Disclosure Statement/PTO-1449 [ ] Copies of IDS Citations
11. [ ] Preliminary Amendment
12. **[X] Return Receipt Postcard**
13. **[X] Small Entity Statement(s)**
14. [ ] Certified Copy of Priority Document(s) (if foreign priority claimed)
15. [ ] Other:____________

UTILITY PATENT APPLICATION TRANSMITTAL *(Only for new nonprovisional application under 37 C.F.R § 1.53(b))*	Docket No. **UDN0003**
	First Inventor or Application Identifier: **Shankar Iyer, et al.**
	Title: **METHOD AND APPARATUS FOR DETERMINING LATENCY BETWEEN MULTIPLE SERVERS AND A CLIENT**
	Express Mail Label No. **EL696097234US**

FEE CALCULATION AND TRANSMITTAL

BASIC FILING FEE *(Small Entity)* Utility Filing Fee				Fee Paid $345
EXTRA CLAIM FEES				
Total Claims [33]	- 20 ** = 15	Extra Claims [0]	X 9	$
Independent Claims [3]	- 3** = 3	Extra Claims [0]	0	
Other Fee:	Recordation			$40
			Total Filing Fee:	**$385**

The Commissioner is hereby authorized to charge and credit Deposit Account No. 501406. A duplicate copy of this sheet is enclosed.
[X] Please credit any overpayment to this account.
[X] Charge any additional filing fees required.

CORRESPONDENCE ADDRESS

24634
PATENT TRADEMARK OFFICE

Customer Number: 24634

Attach Bar Code Label here

Name: **SPEEDERA NETWORKS, INC., ATTENTION: IP-LEGAL DEPT.**

Address: **4800 Great America Parkway, Suite 220**
City: **Santa Clara** State: **CA** ZIP CODE: **95054**

Country: **United States of America** Telephone No. **(408) 970-1500** Fax: **(408) 988-7765**

Name: **Kirk D. Wong, Registration No. 43,284**

Signature: ____________ Date: 9/7/00

STATEMENT CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) & 1.27(f))—INDEPENDENT INVENTOR	DOCKET NO. UDN0003

Applicant(s): **SHANKAR IYER and SRIDHARA LANKA**
Application No.: **To be Assigned**
Filed: **Herewith**

Title: **METHOD AND APPARATUS FOR DETERMINING LATENCY BETWEEN MULTIPLE SERVERS AND A CLIENT**

As a below named inventor, I hereby state that I qualify as an independent inventor as defined in 37 CFR 1.9(c) for purposes of paying reduced fees to the Patent and Trademark Office described herein:

[X] **the specification filed herewith with the title as listed above.**
[ ] the application identified above.
[ ] the patent identified above.

I have not assigned, granted, conveyed, or licensed, and am under no obligation under contract or law to assign, grant, convey, or license, any rights in the invention to any person who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person had made the invention, or to any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern, or organization to which I have assigned, granted, conveyed, or licensed or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

[ ] No such person, concern, or organization exists.
[X] Each such person, concern, or organization is listed below.

SPEEDERA NETWORKS, INC.
Address: 4800 Great America Parkway, Suite 220, Santa Clara, California 95054

Separate statements are required from each named person, concern, or organization having rights to the invention stating their status as small entities (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

Shankar Iyer
Shankar Iyer Date: 8/25/2000

Sridhara
Sridhara Lanka Date: 8·25·2000

STATEMENT CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) & 1.27(c))—SMALL BUSINESS CONCERN	DOCKET NO. UDN0003

Applicant(s): **Shankar Iyer and Sridhara Lanka**
Application No.: **To Be Assigned**
Filed: **Herewith**
Title: **METHOD AND APPARATUS FOR DETERMINING LATENCY BETWEEN MULTIPLE SERVERS AND A CLIENT**

I hereby state that I am

[ ] the owner of the small business concern identified below:
[X] an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF SMALL BUSINESS CONCERN: **SPEEDERA NETWORKS, INC.**
ADDRESS OF SMALL BUSINESS CONCERN: 4800 Great America Parkway, Suite 220, Santa Clara, CA 95054

I hereby state that the above identified small business concern qualifies as a small business concern as defined in 13 CFR Part 121 for purposes of paying reduced fees to the Patent and Trademark Office.

I hereby state that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention described in:

[X] the specification filed herewith with the title as listed above.
[ ] the application identified above.
[ ] the patent identified above.

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights in the invention must file separate statements as to their status as small entities, and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent invention under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern, or organization having rights in the invention is listed below:
[X] No such person, concern, or organization exists.
[ ] Each such person, concern, or organization is listed below.

Separate statements are required from each named person, concern, or organization having rights to the invention stating their status as small entities (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

NAME OF PERSON SIGNING: **Ajit Gupta**
TITLE: Chief Executive Officer and President
ADDRESS OF PERSON SIGNING: 4800 Great America Parkway, Suite 220, Santa Clara, CA 95054

Signature ____________ Date Aug 28, 2000

Method and Apparatus for Determining Latency Between Multiple Servers and a Client

CROSS-REFERENCES TO RELATED APPLICATIONS

The present application claims priority to U.S. Provisional Application No. 60/193,988 filed March 31, 2000 (Attorney Docket No. 4878-US), commonly owned, and hereby incorporated by reference for all purposes.

BACKGROUND OF THE INVENTION

TECHNICAL FIELD

The invention relates to the routing of requests to a networked server in a computer environment. More particularly, the invention relates to determining a dynamic hop count between two nodes across a network in a computer environment.

DESCRIPTION OF THE PRIOR ART

The Internet is a confederation of loosely connected networks that connect computers all around the world. The proliferation of information and immense user population has made the management of information critical. Information is mirrored on multiple servers in order to improve performance, availability, and scalability.

To facilitate the mirroring of information, clients requesting information need to be routed to the optimal server. The traffic management industry routes traffic to the optimal server using different latency metrics including the Round Trip Time (RTT) and hop count. The Round Trip Time measures the time it takes for a packet to travel between a server and a client or another server.

The dynamic hop count between a client and a server is typically derived from the Border Gateway Protocol (BGP). BGP is specified in RFC 1267 published in October 1991 and RFC 1654 published in July 1994. BGP gives the hops between different Autonomous System Numbers (ASN). The dynamic hop count can be used to differentiate latency metrics that might be close in terms of RTT.

The primary function of a BGP speaking system is to exchange network reachability information with other BGP systems. The network reachability information includes information on the Autonomous Systems (AS) that the reachability information traverses. A BGP speaker advertises to its peers, *i.e.*, other BGP speakers that it communicates with, in neighboring ASs only those routes that it uses. The information is sufficient to construct a graph of AS connectivity from which routing loops may be pruned.

Determining the dynamic hop count is currently problematic due to the following reasons:

1. **BGP hops**: The actual number of hops will typically be larger than that advertised by the BGP protocol since BGP only gives the hops between the ASNs.
2. **IP hops**: Getting hop counts from the Internet Protocol (IP) is difficult because the Time To Live (TTL) field that is part of the IP header is not initialized to standard values by the various TCP/IP stack software running on various Operating Systems, *e.g.,* Windows 98, Linux, NT, etc.

The invention described in this patent provides a different and more precise method of determining the dynamic hop count.

It would be advantageous to provide a method and apparatus for determining latency between multiple servers and a client that provides a more precise method of determining dynamic hop counts. It would further be advantageous to provide a method and apparatus for determining latency between multiple servers and a client that reduces the traffic required to measure the hops across the network.

SUMMARY OF THE INVENTION

The invention provides a method and apparatus for determining latency between multiple servers and a client. The system provides a more precise method of determining dynamic hop counts and optimal content servers. In addition, the invention reduces network traffic required for measuring the dynamic hop counts.

A preferred embodiment of the invention receives requests for content server addresses from local domain names servers (LDNS). POPs that can serve the content are determined and sent latency metric requests.

The content server receives the request for latency metrics and looks up the latency metric for the client of the requesting LDNS. Periodic latency probes are sent to the IP addresses in a Latency Management Table. The IP addresses of clients are masked so the latency probes are sent to higher level servers to reduce traffic across the network. The hop count and latency data in the packets sent in response to the latency probes are stored in the Latency Management Table.

The information in the Latency Management Table is used to determine the latency metric from the resident POP to the requesting client before sending the latency metric to the requesting server. The BGP hop count in the Latency Management Table is used for the latency metric upon the first request for an IP address. The latency metric is calculated for subsequent requests of IP addresses using the hop count and RTT data in the Latency Management Table.

Latency metrics from POPs are collected and the inverse relationship of the hop counts in a weighted combination with the RTT are used to determine which latency metric indicates the optimal POP. The address of the optimal POP is then sent to the requesting LDNS.

Other aspects and advantages of the invention will become apparent from the following detailed description in combination with the accompanying drawings, illustrating, by way of example, the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block schematic diagram of a preferred embodiment of the invention measuring latency between servers and a client according to the invention;

Fig. 2 is a block schematic diagram of an example of a preferred embodiment of the invention measuring dynamic hop counts from two POPs to a Border Gateway server according to the invention;

Fig. 3 is a diagram of a Latency Management table according to the invention;

Fig. 4 is a block schematic diagram of an example of differing TTL values in IP packets according to the invention;

Fig. 5 is a block schematic diagram of the inverse relationship of dynamic hop counts in a preferred embodiment of the invention according to the invention; and

Fig. 6 is a block schematic diagram of a task-level viewpoint of a preferred embodiment of the invention according to the invention.

DETAILED DESCRIPTION OF THE INVENTION

The invention is embodied in a method and apparatus for determining latency between multiple servers and a client in a computer environment. A system according to the invention provides a more precise method of determining dynamic hop counts and optimal content servers. In addition, the invention provides a system that reduces network traffic required for measuring the dynamic hop counts.

The invention provides a new method to determine the dynamic hop count between two nodes (client and server). This new method provides a dynamic hop count that is more precise than the hop count obtained using the Border Gateway Protocol (BGP).

The latency probes in a Speedera Network are responsible for determining the latency between the Speedera servers and a client. This latency information is used by the Speedera Domain Name Server (SPDNS) to direct a client to the server that is "closest" to the client in latency.

Referring to Fig. 1, each Speedera Point of Presence (POP) 101, 102 has a probe server 104, 106 and other infrastructure servers 103, 105. When the Client 109 receives a request, it tries to resolve the request from the Local Domain Name Server (LDNS) 108. If the LDNS 108 cannot resolve the request,

it forwards the request to the Speedera Domain Name Server (SPDNS) 107. The SPDNS 107 then routes the request to the optimal POP 101, 102 by determining the server "closest" to the LDNS 108.

In this example, the Client 109 performs a name lookup for web content images.rich.com/images 110. The local DNS (LDNS) client 108 forwards the request 111 to the SPDNS 107.

The SPDNS 107 requests latency information 112, 113 from the Speedera probes 104, 106 at locations that can serve images.rich.com (POP1 101 and POP2 102). The latency probes 104, 106 initiate probes 117, 118 to the LDNS 108. The latency probes 104, 106 return the latency metrics 115, 116 to the SPDNS 107.

The main components that are used by the invention to determine latency metrics are as follows:

- RTT time from the latency probe to LDNS.
- ASN (Autonomous System Number) routing information derived from Border Gateway Protocol (BGP).
- Dynamic hop count from the latency probe to LDNS.

BGP is a standard algorithm implemented in routers. It is a routing protocol that is used between large routers covering large administrative domains. All routes that are available within a network are exported to another network in an abbreviated form. The network reachability information includes information on the Autonomous Systems (AS) that the reachability information traverses. A BGP speaker (router) advertises to its peers, *i.e.*, other BGP speakers that it communicates with, in neighboring ASs only those routes that it uses.

An AS is a set of routers under a single technical administration, using an interior gateway protocol and common metrics to route packets within the AS, and using an exterior gateway protocol to route packets to other ASs.

The latency probe uses a UDP Reverse Name Lookup and Traceroute to determine RTT and dynamic hop count. Reverse Name Lookup is a standard DNS query that specifies a client IP address and asks for the client name. Traceroute is a specific format of a packet that is sent between routers to indicate if a packet has reached a destination.

Obtaining absolute hop counts is ideal, however, this is not a realistic goal in the real world. Quite often some of the local DNSs are unavailable because they are sitting behind a firewall or the DNS will drop the packet thinking that it is a denial of service attack. With these restrictions, it is not possible to reliably use an absolute hop count. Additionally, as described below, the IP packet information is not uniform throughout the network.

With respect to Fig. 2, most of the time the ASN 203 is easily reached. Even from various diverse points it is possible to converge onto the AS 203. However, it is the path required to reach the AS 203 from each server that is the important factor.

POP1 201 and Pop2 202 must each find the distance from themselves to the ASN 203. The latency for each path must also be found. The distance and latency are calculated by using the hop count and latency time, respectively.

Currently, it is only possible to measure hop count and latency time by sending a packet to the destination. However, as previously mentioned, most of the time the packet is sent back as not being able to reach the destination.

The invention aggregates the client 208 and the DNS 207 and assumes that they are co-located. Additionally, the latency and the hop counts are measured up to the Border Gateway (BG) 206. Once the autonomous system is entered, the hop counts are not as important. The distance from the BG 206 to the client 208 is the same from either POP 201, 202 at that point. Therefore, the relevant distance is to the BG 206. In other words, the distance T1 210 and T2 211 are most likely not equal, but the distance T3 212 is the same for both POPs 201, 202.

The address of the client 208 is masked to the IP prefix of the ASN 203. For example, if the address of the client 208 is 4.10.20.30 then the address is masked to the ASN 203 which is 4.0.0.0. The mask can vary depending on the granularity desired. For example, the first eight bits may be masked to the DNS, *i.e.*, the address would then be masked to 4.10.20.00. This can be adjusted depending on the size of the network.

The data in the BGP table does not change over a period of time. Incremental updates are sent as the routing tables change. However, BGP does not require periodic refresh of the entire BGP routing table. Therefore, a BGP speaker must retain the current version of the entire BGP routing tables of all of its peers for the duration of the connection.

The invention performs its measurements a few times a day to achieve a good picture of the network status. The invention reduces Internet traffic by aggregating at a higher level and therefore requires less probes.

Referring to Fig. 3, a latency management table is used by the invention. The table contains the following fields: IP address 301; BGP hop count 302; and Trace data 303. The IP address field 301 contains the IP addresses that the server is responsible for. BGP hop counts 302 is taken from the BGP routing

table for the particular IP address. The Trace data field 303 contains the latency and hop count information obtained by the latency probes.

The BGP hop count 302 is used by the SPDNS when the first request from an LDNS comes in for a particular IP address. No previous connection has been established at this point. This is because the dynamic hop count to the BG takes some time to actually measure. Subsequent requests use the actual BG hop count measured by the system located in the Trace field 303. The latency measurement is also taken and the combination of the hop count and latency measurement in the Trace field 303 is used to determine which SPDNS is closest to the client.

The invention does not need to get absolute hop counts from the various probe servers to the LDNS. It is sufficient to determine the relative hop counts between the probe servers and the LDNS and use this information in arriving at relative latency metrics between the client and the probe servers. This information can then be used by the Speedera DNS to choose between the service points and locate a server that is closest to the client.

With respect to Fig. 4, the invention is based on the very safe assumption that the target LDNS 402 will always send out packets with a fixed TTL independent of the SPDNS 401 that initiated a probe. Getting hop counts from the Internet Protocol (IP) is difficult because the Time To Live (TTL) field that is part of the IP header is not initialized to standard values by the various TCP/IP stack software running on various Operating Systems, *e.g.,* Windows 98, Linux, NT, etc.

In this example, the SPDNS 401 originates a packet with a TTL of 64 403. The LDNS 402 always sends out packets with its own independent TTL value of 128 404. DNS's will send whatever TTL value that they prefer. The values may differ between each DNS, but each DNS is consistent in its TTL value.

Referring to Fig. 5, the two probe servers (Probe1 502 and Probe2 503) have initiated probes to the LDNS 501. The LDNS 501 returns a response to each of the probe servers. Here, the LDNS 501 sets the TTL in the response IP packet to *R* 504, 505. The TTL will be decremented by 1 each time the packet passes through a router (hop). In this example, the packets go through *H1* hops 506 between LDNS 501 and Probe1 502 and *H2* hops 507 between LDNS 501 and Probe2 503. The TTL of the response packet that arrives at Probe1 502 will be *(R-H1)* and which arrives at Probe2 503 will be *(R-H2)*. The TTL of the response at Probe1 502 and Probe2 503 will be in inverse relation to the number of hops 508, 509 (The fewer the number of hops, the higher the TTL). Thus, a relative hop count can be arrived at by subtracting the TTL in the response packet from a fixed value.

The latency metric is a weighted combination of the RTT and the hop count. The latency metric is used to determine the server that is the most efficient for accessing a client. The invention precisely determines the hop count metric between client and server.

The dynamic hop count metric derived through the invention is more accurate than the hop count derived from BGP. As a result, requests will be routed more accurately to the optimal server resulting in improved performance. The invention improves performance in multiple Internet infrastructure products including performance monitoring, WAN traffic management, and content distribution.

With respect to Fig. 6, a task level viewpoint of the invention is shown. Requests for content server addresses from LDNS's are received by the Receive IP Address Request module 605. The Receive IP Address Request module 605 sends the content request to the Request Latency Metrics module 606. POPs that can serve the content are retrieved from the Server Table 609 by the Request Latency Metrics module 606. The Request Latency Metrics module 606

then sends latency metric requests to the POPs that can serve the content and notifies the Determine Optimal Server module 608 that latency metrics are expected.

Request for latency metrics activate the Send Latency Metric module 601 to lookup the latency metric for the requesting LDNS's client. The Send Latency Probe module 603 sends latency probes to the IP addresses in the Latency Management Table 604. The IP addresses of clients are masked so the latency probes are sent to higher level servers as detailed above. Packets sent in response to the latency probes sent by the Send Latency Probe module 603 are received by the Receive Response Packet module 602. Hop count and latency data are stored into the Latency Management Table 604.

The Send Latency Metric module 601 uses the information in the Latency Management Table 604 to determine the latency metric from the resident POP to the requesting LDNS's client before sending the latency metric to the requesting server. The Send Latency Metric module 601 uses the BGP hop count in the Latency Management Table 604 for its calculations upon the first request for an IP address. The latency metric is calculated for subsequent requests of IP addresses by the Send Latency Metric module 601 using the hop count and RTT data obtained from the Receive Response Packet module 602.

Latency metrics from POPs are received by the Receive Latency Metrics module 607. The latency metrics are sent to the Determine Optimal Server module 608. The Determine Optimal Server module 608 gathers the expected latency metrics and uses the inverse relationship of the hop counts in a weighted combination with the RTT to determine which latency metric indicates the optimal POP. The Determine Optimal Server module 608 then sends the address of the optimal POP to the requesting LDNS.

Although the invention is described herein with reference to the preferred embodiment, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of the present invention. Accordingly, the invention should only be limited by the Claims included below.

CLAIMS

1. A process for determining latency between multiple servers and a client across a network in a computer environment, comprising the steps of:

receiving a request for latency metrics on a content server;

wherein said latency metric request specifies a particular client;

providing a latency management table;

wherein said latency management table comprises a list of IP addresses along with corresponding Border Gateway Protocol (BGP) hop counts, dynamic hop counts, and Round Trip Times (RTT);

looking up the latency metric for said client in said latency management table;

sending said latency metric to the requesting server;

wherein the BGP hop count for said client in said latency management table is used for said latency metric upon the first request for said client; and

wherein the dynamic hop count and RTT data for said client in said latency management table are used for said latency metric for subsequent requests for said client.

2. The process of Claim 1, further comprising the steps of:

sending periodic latency probes to the IP addresses in said latency management table;

receiving response packets for said latency probes; and

recording the dynamic hop count and latency (RTT) data in said latency management table.

3. The process of Claim 2, wherein periodic latency probes are sent to a higher level server of a client by masking said client's IP address in said latency management table.

4. The process of Claim 1, further comprising the steps of:

receiving requests for a content server address from said client;

sending a latency metric request to the appropriate content servers;

receiving latency metric data from said content servers;

determining the optimal content server for said client; and

sending said optimal content server's address to said client.

5. The process of Claim 4, wherein said determining step gathers the expected latency metrics and uses the inverse relationship of the hop counts in said latency metric data in a weighted combination with the RTT in said latency metric data to determine which latency metric data indicates the optimal content server.

6. An apparatus for determining latency between multiple servers and a client across a network in a computer environment, comprising:

a module for receiving a request for latency metrics on a content server;

wherein said latency metric request specifies a particular client;

a latency management table;

wherein said latency management table comprises a list of IP addresses along with corresponding Border Gateway Protocol (BGP) hop counts, dynamic hop counts, and Round Trip Times (RTT);

a module for looking up the latency metric for said client in said latency management table;

a module for sending said latency metric to the requesting server;

wherein the BGP hop count for said client in said latency management table is used for said latency metric upon the first request for said client; and

wherein the dynamic hop count and RTT data for said client in said latency management table are used for said latency metric for subsequent requests for said client.

7. The apparatus of Claim 6, further comprising:

a module for sending periodic latency probes to the IP addresses in said latency management table;

a module for receiving response packets for said latency probes; and

a module for recording the dynamic hop count and latency (RTT) data in said latency management table.

8. The apparatus of Claim 7, wherein periodic latency probes are sent to a higher level server of a client by masking said client's IP address in said latency management table.

9. The apparatus of Claim 7, further comprising:

a module for receiving requests for a content server address from said client;

a module for sending a latency metric request to the appropriate content servers;

a module for receiving latency metric data from said content servers;

a module for determining the optimal content server for said client; and

a module for sending said optimal content server's address to said client.

10. The apparatus of Claim 9, wherein said determining module gathers the expected latency metrics and uses the inverse relationship of the hop counts in said latency metric data in a weighted combination with the RTT in said latency metric data to determine which latency metric data indicates the optimal content server.

11. A program storage medium readable by a computer, tangibly embodying a program of instructions executable by the computer to perform method steps for determining latency between multiple servers and a client across a network in a computer environment, comprising the steps of:

receiving a request for latency metrics on a content server;

wherein said latency metric request specifies a particular client;

providing a latency management table;

wherein said latency management table comprises a list of IP addresses along with corresponding Border Gateway Protocol (BGP) hop counts, dynamic hop counts, and Round Trip Times (RTT);

looking up the latency metric for said client in said latency management table;

sending said latency metric to the requesting server;

wherein the BGP hop count for said client in said latency management table is used for said latency metric upon the first request for said client; and

wherein the dynamic hop count and RTT data for said client in said latency management table are used for said latency metric for subsequent requests for said client.

12. The method of Claim 11, further comprising the steps of:

sending periodic latency probes to the IP addresses in said latency management table;

receiving response packets for said latency probes; and

recording the dynamic hop count and latency (RTT) data in said latency management table.

13. The method of Claim 12, wherein periodic latency probes are sent to a higher level server of a client by masking said client's IP address in said latency management table.

14. The method of Claim 11, further comprising the steps of:
receiving requests for a content server address from said client;
sending a latency metric request to the appropriate content servers;
receiving latency metric data from said content servers;
determining the optimal content server for said client; and
sending said optimal content server's address to said client.

15. The method of Claim 14, wherein said determining step gathers the expected latency metrics and uses the inverse relationship of the hop counts in said latency metric data in a weighted combination with the RTT in said latency metric data to determine which latency metric data indicates the optimal content server.

Method and Apparatus for Determining Latency Between Multiple Servers and a Client

ABSTRACT

A method and apparatus for determining latency between multiple servers and a client receives requests for content server addresses from local domain names servers (LDNS). POPs that can serve the content are determined and sent latency metric requests. The content server receives the request for latency metrics and looks up the latency metric for the requesting client. Periodic latency probes are sent to the IP addresses in a Latency Management Table. The IP addresses of clients are masked so the latency probes are sent to higher level servers to reduce traffic across the network. The hop count and latency data in the packets sent in response to the latency probes are stored in the Latency Management Table and is used to determine the latency metric from the resident POP to the requesting client before sending the latency metric to the requesting server. The BGP hop count in the Latency Management Table is used for the latency metric upon the first request for an IP address. The latency metric is calculated for subsequent requests of IP addresses using the hop count and RTT data in the Latency Management Table. Latency metrics from POPs are collected and the inverse relationship of the hop counts in a weighted combination with the RTT are used to determine which latency metric indicates the optimal POP. The address of the optimal POP is then sent to the requesting LDNS.

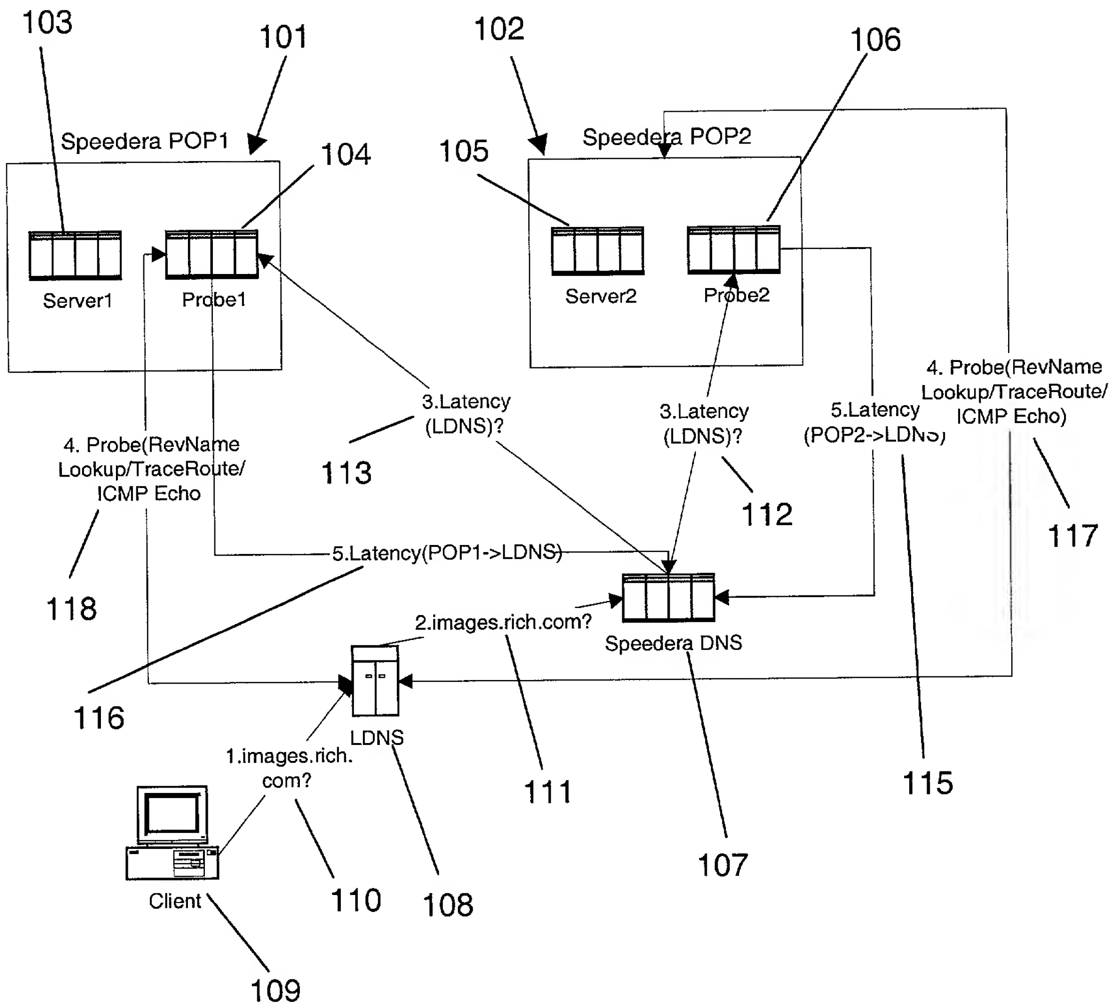
103
101
102
106
Speedera POP1
104
105
Speedera POP2
Server1
Probe1
Server2
Probe2
4. Probe(RevName
Lookup/TraceRoute/
ICMP Echo)
3.Latency
(LDNS)?
3.Latency
(LDNS)?
5.Latency
(POP2->LDN
4. Probe(RevName
Lookup/TraceRoute/
ICMP Echo
113
112
117
118
5.Latency(POP1->LDNS)
Speedera DNS
2.images.rich.com?
116
LDNS
1.images.rich.
com?
111
115
107
Client
110
108
109

Fig. 1

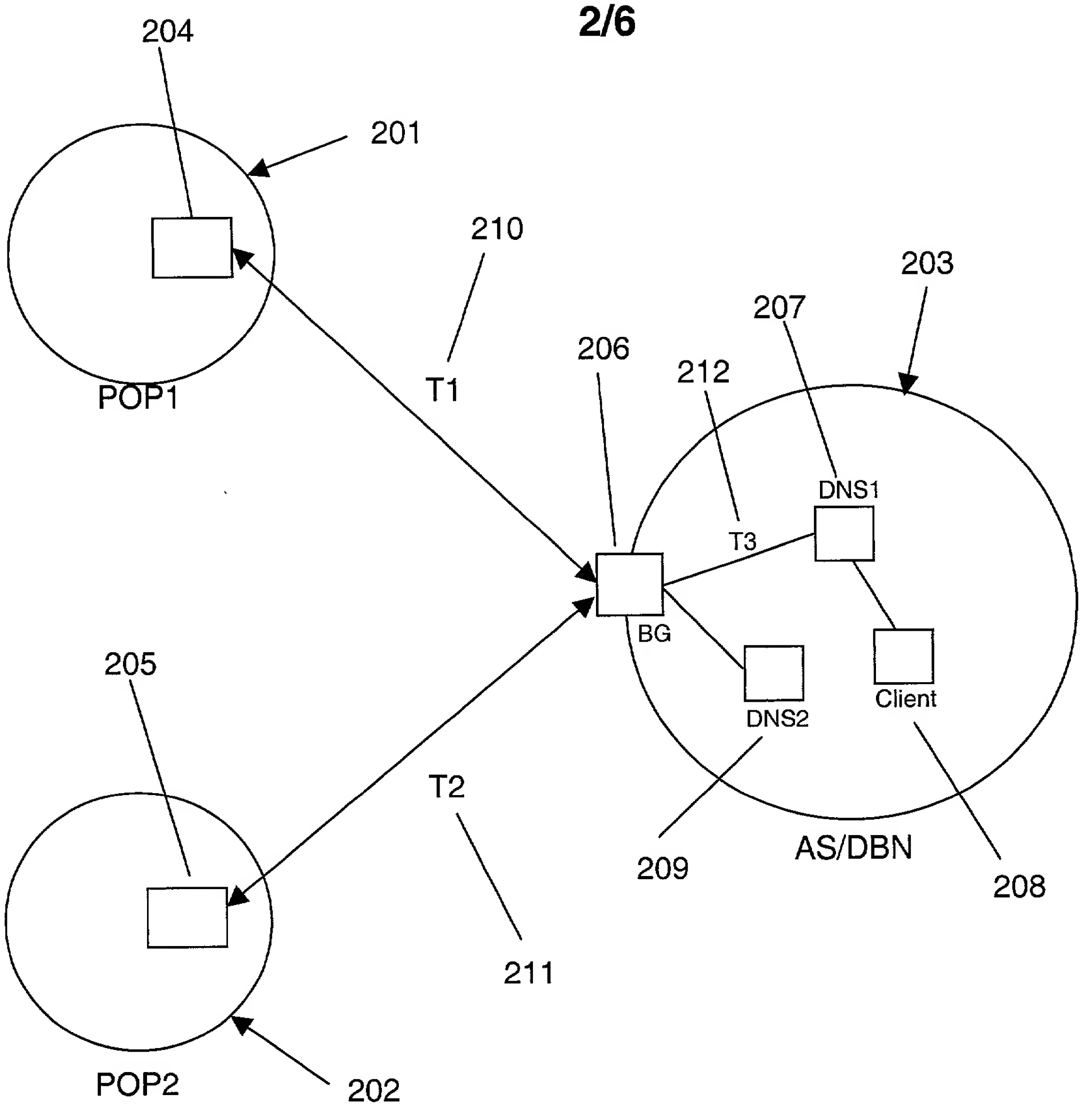
204
201
210
203
207
206
212
POP1
T1
DNS1
T3
BG
205
DNS2
Client
T2
209
AS/DBN
208
211
POP2
202

Fig. 2

301	302	303
IP address	BGP hop count	Trace
4.0.0.0/8	3	10/100ms
10.6.0.0/4	2	8/80ms

Fig. 3

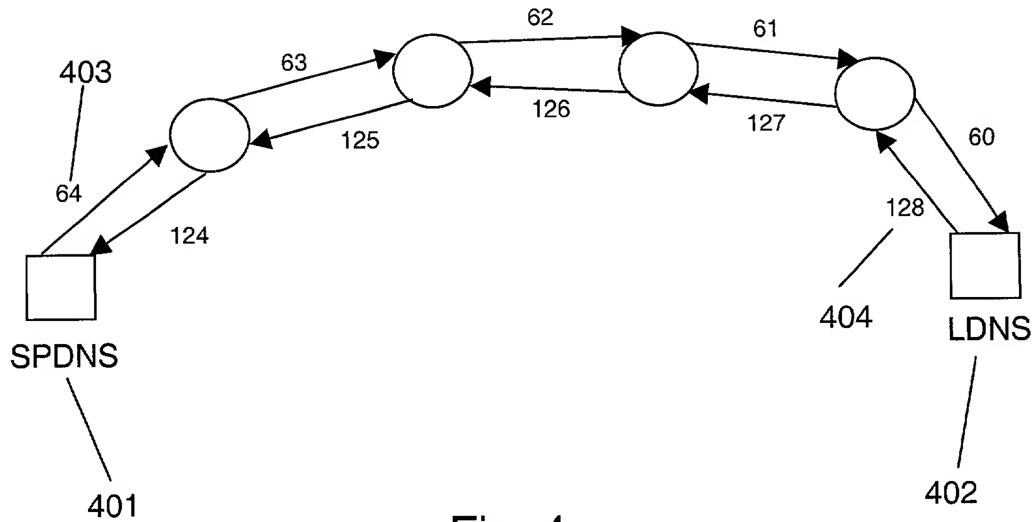

Fig. 4

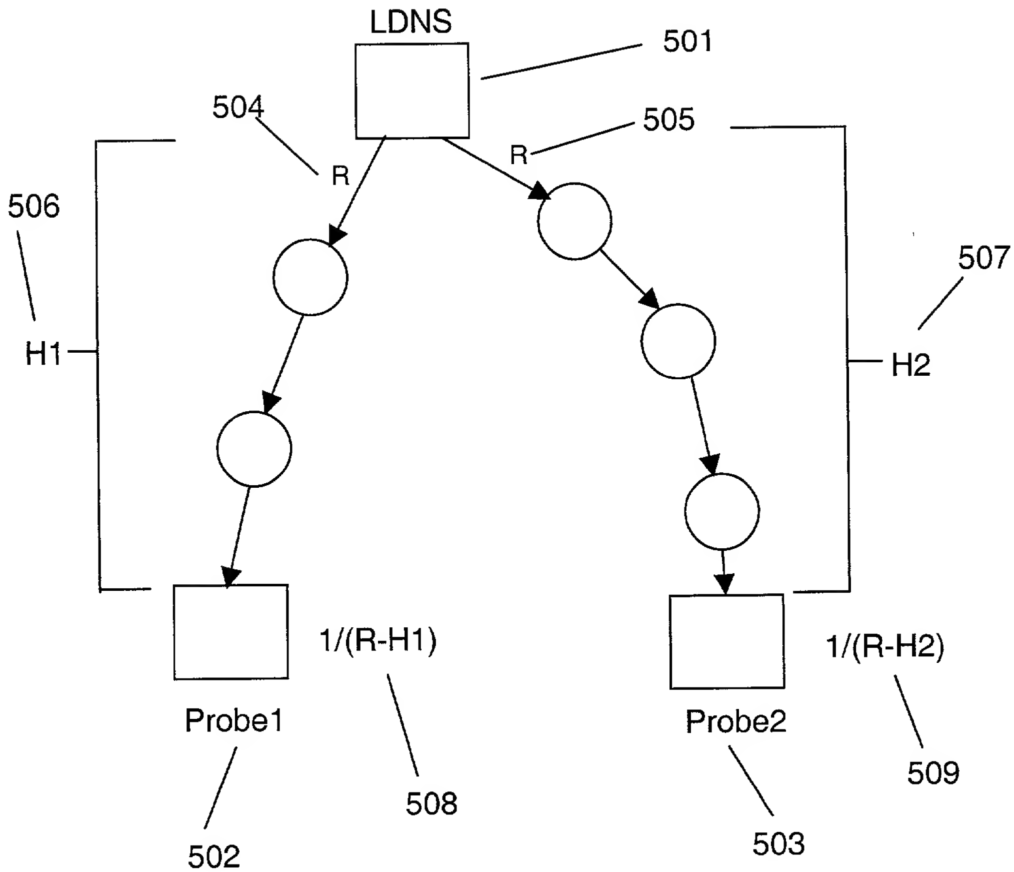
LDNS
501
504
505
R
R
506
507
H1
H2
1/(R-H1)
1/(R-H2)
Probe1
Probe2
502
503
508
509

Fig. 5

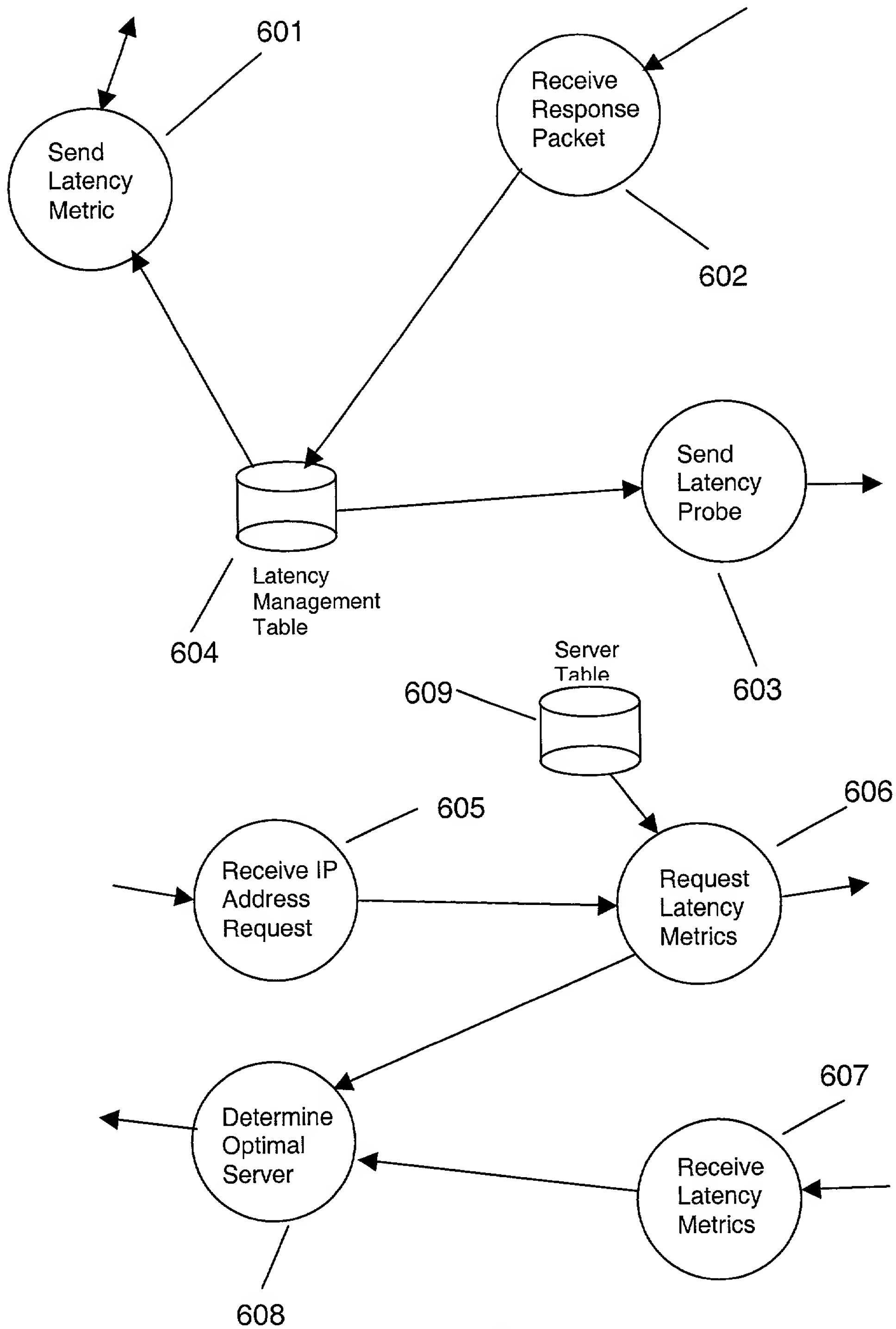
601
Send Latency Metric
Receive Response Packet
602
Send Latency Probe
Latency Management Table
604
603
Server Table
609
605
606
Receive IP Address Request
Request Latency Metrics
607
Determine Optimal Server
Receive Latency Metrics
608

Fig. 6

DECLARATION FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name;

I believe I am the original, first, and sole inventor (if only one name is listed below) or an original, first, and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

METHOD AND APPARATUS FOR DETERMINING LATENCY BETWEEN MULTIPLE SERVERS AND A CLIENT

the specification of which (check one) X **is attached hereto**, or ___ was filed on __________________ as Application Serial No. __________________ and was amended on _______ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, Section 1.56(a).

===

I hereby claim foreign priority benefits under Title 35, United Sates Code, Section 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s)	Priority Claimed Yes	Priority Claimed No
______________________ Number Country Day/Month/Year Filed	____	____
______________________ Number Country Day/Month/Year Filed	____	____

===

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

KIRK D. WONG, REG. NO. 43,284

SEND CORRESPONDENCE TO:

Customer No. 24634
ATTENTION: IP-LEGAL DEPARTMENT
Speedera Networks, Inc.
4800 Great America Parkway, Suite 220
Santa Clara, CA 95054

Speedera File No. UDN0003

I hereby claim the benefit under Title 35, United States code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application or PCT international application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application Ser. No.	Filing Date	Status: Patented, Pending, Abandoned
60/193,988	March 31, 2000	Pending

==

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor: **Shankar Iyer**

Inventor's signature ______________ 8/25/00
Date

Residence 10551 Cypress Court, Cupertino, CA

Post Office Address Same

Citizenship INDIA

Full name of second inventor: **Sridhara Lanka**

Inventor's signature ______________ 8-25-2000
Date

Residence 1101 Littleoak Cr, San Jose CA-95129

Post Office Address Same

Citizenship India